2011 American Control Conference
on O'Farrell Street, San Francisco, CA, USA
June 29 - July 01, 2011

Multi-agent Systems Reaching Optimal Consensus with Directed Communication Graphs

Guodong Shi, Karl Henrik Johansson and Yiguang Hong

***Abstract*— In this paper, we investigate an optimal consensus problem for multi-agent systems with directed interconnection topologies. Based on a nonlinear distributed coordination rule with switching directed communicating graphs, the considered multi-agent system achieves not only a consensus, but also an optimal one by agreeing within the global solution set of a sum of objective functions corresponding to multiple agents. The optimal solution set convergence and consensus analysis are given respectively with the help of convex analysis and nonsmooth analysis.**

***Index Terms*— Multi-agent systems, consensus, distributed optimization, directed graph**

I. INTRODUCTION

Cooperative control of multi-agent systems becomes an active research area from the beginning of this century, and rapid developments of distributed control protocols via interconnected communication have been made to achieve the collective tasks (referring to [20], [17], [13], [23], [10], [8], [9], [18]).

Consensus and formation are important problems of multi-agent coordination, since in reality it is usually required that all the agents (such as robots or vehicles) achieve the desired relative position and the same velocity. Connectivity plays a key role in the coordination of multi-agent network, and various connectivity conditions to describe frequently switching topologies in different cases. The "joint connection" or similar concepts are important in the analysis of stability and convergence to guarantee multi-agent coordination. Uniformly jointly-connected conditions have been employed for different problems ([20], [17], [21], [6]). On the other hand, $[t, \infty)$-joint connectedness is the most general form to secure the global coordination, which is also proved to be necessary in many situations ([23], [18]).

Moreover, multi-agent optimization has attracted much attention in recent years(referring to [29], [30], [25]). In [29], a distributed algorithm which solves a special class of optimization problems by using only peer-to-peer communication was proposed. In [30], a subgradient method in combination with a consensus process was given for solving coupled optimization problems in a distributed way with fixed undirected graph. Then in [27], the authors showed the convergence bound for sub-gradient based multi-agent optimization in various connectivity assumptions with time-varying graphs. In [28], a constrained consensus problem for multi-agent networks is considered when each agent is restricted to lie in its own convex set. However, in most existing works, the optimization model was assumed to be a convex optimization problem, and convergence to the optimal solution set was usually missing. Moreover, the mostly considered multi-agent model in existing works were with discrete-time dynamics.

The objective of this paper is to study the distributed optimization of multi-agent systems with directed communication graphs. In other words, we aim to provide the optimal consensus protocols of the multi-agent systems with switching communication topologies. Different from the existing results, we obtain a global consensus and convergence to the optimal solution set of the coupled objective function which is a sum of objective functions corresponding to multiple agents.

The paper is organized as follows. In Section 2, necessary preliminaries and problem formulation are given. In Section 3, the main result is proposed on optimal consensus, and then discuss the distance function estimation for further analysis. Then, in Section 4, the optimal solution set convergence analysis is carried out, based on which we give the proof the main result of the paper. Finally, in Section 5 concluding remarks are given.

II. PROBLEM FORMULATION

In this section, we formulate our problem and introduce related preliminary knowledge.

Consider a multi-agent system with agent set $\mathcal{V} = \{1, 2, \cdots, N\}$, for which the dynamics of each agent is the following first-order integrator:

$$\dot{x}_i = u_i, \quad i = 1, \cdots, N \tag{1}$$

where $x_i \in R^m$ represents the state of agent i, and u_i is the control input. The agent can be viewed as a node in a graph.

The control objective is to reach a consensus for this group of autonomous agents, and meanwhile to cooperatively solve the following optimization problem

$$\min_{z \in R^m} \quad F(z) = \sum_{i=1}^{N} f_i(z) \tag{2}$$

where $f_i : R^m \to R$ represents the cost function of agent i, observed by agent i only, and z is a decision vector.

This work has been supported in part by the Knut and Alice Wallenberg Foundation, the Swedish Research Council and the NNSF of China under Grants 60874018 and 60821091.

G. Shi and K. Johansson are with ACCESS Linnaeus Centre, School of Electrical Engineering, Royal Institute of Technology, Stockholm 10044, Sweden.Email: guodongs@kth.se, Kallej@ee.kth.se

Y. Hong is with Key Laboratory of Systems and Control, Institute of Systems Science, Chinese Academy of Sciences, Beijing 100190, China. Email: yghong@iss.ac.cn

978-1-4577-0079-8/11/$26.00 ©2011 AACC

Denote the global optimal solution set (suppose it exists) of function f_i by $\mathcal{S}_i$, i.e.,

$$\mathcal{S}_i \doteq \{y \mid f_i(y) = \min_{z\in R^m} f_i(z)\}, \quad i = 1, \cdots, N.$$

A set $K \subset R^d$ is said to be convex if $(1-\alpha)x + \alpha y \in K$ whenever $x \in K, y \in K$ and $0 \le \alpha \le 1$. An assumption for each $\mathcal{S}_i$ is stated in the following:

Assumption 1. $\mathcal{S}_i$ is convex for $i = 1, \cdots, N$, and $\bigcap_{i=1}^{N} \mathcal{S}_i \neq \emptyset$.

Remark 2.1: Note that a function $f : R^m \to R$ is said to be convex if it satisfies [24]

$$f(\alpha v + (1-\alpha)w) \le \alpha f(v) + (1-\alpha) f(w), \tag{3}$$

for all $v, w \in R^m$ and $0 \le \alpha \le 1$. Moreover, if the cost function f_i is a convex function for $i = 1, \cdots, N$, optimization problem (2) is $v, w \in \mathcal{S}_i$ a convex optimization problem since $F(x)$ is then convex in this case. However, when f_i is convex, we have that, for any $v, w \in \mathcal{S}_i$ and $0 \le \alpha \le 1$,

$$\begin{aligned} \min_{z\in R^m} f_i(z) &\le f_i(\alpha v + (1-\alpha)w) \\ &\le \alpha f_i(v) + (1-\alpha) f_i(w) \\ &= \min_{z\in R^m} f_i(z). \end{aligned} \tag{4}$$

This implies that $\alpha v + (1-\alpha)w \in \mathcal{S}_i, 0 \le \alpha \le 1$, which leads to that $\mathcal{S}_i$ is a convex set. On the other hand, there are many cases that $\mathcal{S}_i$ is a convex set while f_i is not a convex function. Therefore, in this sense to assume that $\mathcal{S}_i$ is a convex set is more generalized than that f_i is a convex function.

Denote the global optimal solution set of cost function $F(x)$ by $\mathcal{S}_0$, i.e., $\mathcal{S}_0 \doteq \{y \mid F(y) = \min_{z\in R^m} F(z)\}$. Then with Assumption 1, it is obvious to see $\mathcal{S}_0 = \bigcap_{i=1}^{N} \mathcal{S}_i$.

A. Communication Network Model

In this subsection, let us describe the communication rule, i.e., the information exchange model for the considered multi-agent network.

First we will introduce some concepts in graph theory (referring to [3] for details). A directed graph (digraph) $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consists of a finite set $\mathcal{V}$ of nodes and an arc set $\mathcal{E}$, in which an arc is an ordered pair of distinct nodes of $\mathcal{V}$. $(i, j) \in \mathcal{E}$ describes an arc which leaves i and enters j. A *walk* in digraph $\mathcal{G}$ is an alternating sequence $\mathcal{W} : i_1 e_1 i_2 e_2 \cdots e_{m-1} i_m$ of nodes i_κ and arcs $e_\kappa = (i_\kappa, i_{\kappa+1}) \in \mathcal{E}$ for $\kappa = 1, 2, \cdots, m-1$. A walk is called a *path* if the nodes of this walk are distinct, and a path from i to j is denoted as $\widehat{(i,j)}$. $\mathcal{G}$ is said to be *strongly connected* if it contains path $\widehat{(i,j)}$ and $\widehat{(j,i)}$ for every pair of nodes i and j.

In this paper, the communication in the multi-agent network is supposed to be directed and time-varying. The system topology is modeled as a time-varying directed graph $\mathcal{G}_{\sigma(t)} = (\mathcal{V}, \mathcal{E}_{\sigma(t)})$, where $\mathcal{E}_{\sigma(t)}$ represents the arc (link) set defined by a piecewise constant switching signal function $\sigma : [0, +\infty) \to \mathcal{P}$ with $\mathcal{P}$ as the set of all possible interconnection topologies. At time t, node $i \in \mathcal{V}$ can receive the information from $j \in \mathcal{V}$ if there is an arc $(j, i) \in \mathcal{E}_{\sigma(t)}$ from j to i, and in this way, j is said to be a *neighbor* of agent i. As usual, we assume there is a *dwell time*, denoted by a constant τ_D for $\sigma(t)$, as a lower bound between two switching times.

Denote the joint digraph of $\mathcal{G}_{\sigma(t)}$ in time interval $[t_1, t_2)$ with $t_1 < t_2 \le +\infty$ by

$$\mathcal{G}([t_1, t_2)) = \cup_{t\in[t_1,t_2)} \mathcal{G}(t) = (\mathcal{V}, \cup_{t\in[t_1,t_2)} \mathcal{E}_{\sigma(t)}). \tag{5}$$

Then $\mathcal{G}_{\sigma(t)}$ is said to be *uniformly jointly strongly connected* (UJSC) if There exists a constant $T > 0$ such that $\mathcal{G}([t, t+T))$ is strongly connected for any $t \ge 0$.

B. Distributed Control Law

In this subsection, we introduce the neighbor-based control laws for the agents.

Let K be a closed convex subset in R^d and denote $|x|_K \triangleq \inf\{|x - y| \mid y \in K\}$, where $|\cdot|$ denotes the Euclidean norm for a vector or the absolute value of a scalar. Then we can associate to any $x \in R^d$ a unique element $\mathcal{P}_K(x) \in K$ satisfying $|x - \mathcal{P}_K(x)| = |x|_K$, where the map $\mathcal{P}_K$ is called the projector onto K and

$$\langle \mathcal{P}_K(x) - x, \mathcal{P}_K(x) - y \rangle \le 0, \quad \forall y \in K. \tag{6}$$

Clearly, $|x|_K^2$ is continuously differentiable at point x, and (see [1])

$$\nabla |x|_K^2 = 2(x - \mathcal{P}_K(x)). \tag{7}$$

Denote $x = (x_1, \cdots, x_N)^T \in R^{Nm}$ and let continuous function $a_{ij}(x, t) > 0$ be the weight of arc (j, i), for $i, j \in \mathcal{V}$. Let $N_i(\sigma(t))$ represent the set of agent i's neighbors. Then we present the control law for the agents:

$$u_i = \sum_{j\in N_i(\sigma(t))} a_{ij}(x, t)(x_j - x_i) + \mathcal{P}_{\mathcal{S}_i}(x_i) - x_i. \tag{8}$$

Remark 2.2: In practice, the weights for a multi-agent network, a_{ij}, may not be constant because of the complex communication and environment uncertainties, and then the multi-agent system become time-varying or nonlinear (referring to [21], [18], [23]). Here $a_{ij}(x, t)$ is written in a general form simply for convenience, and global information is not required in the study. For example, a_{ij} can depend only on the state of x_i, time t and x_j ($j \in N_i$), which is certainly a special form of $a_{ij}(x, t)$. In this case, the control laws of form (8) are still decentralized.

Remark 2.3: In (8), we suppose that agent i can observe the vector $\mathcal{P}_{\mathcal{S}_i}(x_i) - x_i$ based on the information of f_i. In practice, $\mathcal{S}_i$ may be solved by agent i beforehand, and then the control is made based on the information of $\mathcal{S}_i$. In some other cases, vector $\mathcal{P}_{\mathcal{S}_i}(x_i) - x_i$ may also be obtained by agent i directly based on the information of f_i. For example, if $f_i = |x_i|_{K_i}^{\lambda}$ for some constant $\lambda > 0$ and convex set K_i, then one has $\mathcal{P}_{\mathcal{S}_i}(x_i) - x_i = \frac{1}{2} \nabla f_i^{\frac{2}{\lambda}}$.

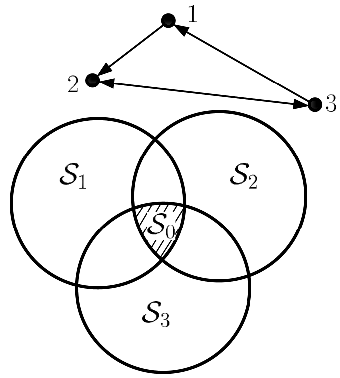

Fig. 1. The goal of the agents is to achieve a consensus in $\mathcal{S}_0$.

Without loss of generality, we assume the initial time $t = 0$, and the initial condition $x^0 = (x_1(0), \cdots, x_n(0))^T \in R^{nd}$. Moreover, for the weights $a_{ij}(x,t)$, we use the following assumption.

Assumption 2. There are $0 < a_* \leq a^*$ such that $a_* \leq a_{ij}(x,t) \leq a^*,\ x \in R^{Nm}, t \geq 0$.

With (1) and (8), the closed loop system is expressed by

$$\dot{x}_i = \sum_{j \in N_i(\sigma(t))} a_{ij}(x,t)(x_j - x_i) + \mathcal{P}_{\mathcal{S}_i}(x_i) - x_i,\ i = 1, \cdots, N. \tag{9}$$

Let $x(t)$ be the trajectory of (9) with initial condition $x(0) = x^0$. Then the considered optimal consensus is defined as following (see Fig. 1).

Definition 2.1: (i) A global *optimal solution set convergence* for System (9) is achieved if

$$\lim_{t \to +\infty} |x_i(t)|_{\mathcal{S}_0} = 0, \quad i = 1, \cdots, N \tag{10}$$

for any initial condition $x^0 \in R^{mN}$.

(ii) A global *consensus* for System (9) is achieved if

$$\lim_{t \to +\infty} x_i(t) - x_j(t) = 0, \quad i, j = 1, \cdots, N \tag{11}$$

for any initial condition $x^0 \in R^{mN}$.

(iii) A global *optimal consensus* is achieved for System (9) if both (i) and (ii) hold.

Remark 2.4: If both (10) and (11) hold, one has

$$\begin{aligned} \lim_{t \to +\infty} \dot{x}_i &= \lim_{t \to +\infty} \sum_{j \in N_i(\sigma(t))} a_{ij}(x,t)(x_j - x_i) \\ &\quad + \mathcal{P}_{\mathcal{S}_i}(x_i) - x_i \\ &= 0. \end{aligned} \tag{12}$$

Thus, it follows that there exists $z_* \in \mathcal{S}_0$ such that $\lim_{t \to +\infty} x_i(t) = z_*$, $i = 1, \cdots, N$.

III. Main Result

In this section, we give the main result and then some basic results for its proof.

The main difficulties to obtain optimal consensus result from the fact that we have to consider the consensus and the convergence to the optimal solution together. Control rule in the form of (8) without the term $\mathcal{P}_{\mathcal{S}_i}(x_i) - x_i$ has been studied for consensus [13], [21], [18]. However, if the agents also try to solve the optimization problem (2) cooperatively, the term like $\mathcal{P}_{\mathcal{S}_i}(x_i) - x_i$ is then inevitable. In fact, the term $\mathcal{P}_{\mathcal{S}_i}(x_i) - x_i$ could coincide the subgradient of f_i in many cases, and then (8) will be consistent with the subgradient method for multi-agent optimization [27], [30]. Therefore, there is usually a tradeoff between consensus and optimization, and it is hard to achieve both of them.

In this paper, we suppose that Assumptions 1 and 2 always hold. The following is the main result of the paper.

Theorem 3.1: System (9) achieves an optimal consensus if $\mathcal{G}_{\sigma(t)}$ is uniformly jointly strongly connected (UJSC).

To prove Theorem 3.1, on one hand, we have to prove all the agents converge to the global optimal solution set $\mathcal{S}_0$, and on the other hand we have to verify that a consensus is also achieved.

To do this, we will first show a method to estimate the distance function.

Define $d_i(t) = |x_i(t)|^2_{\mathcal{S}_0}$ and let

$$\bar{d}(t) = \max_{i \in \mathcal{V}} d_i(t)$$

be the maximum among all the agents.

According to the definition of $\bar{d}(t)$, it is easy to see that usually it is not continuously differentiable. However, $\bar{d}(t)$ is indeed locally Lipschitz. Thus, we can still analyze the Dini derivative of $\bar{d}(t)$ to study its convergence property.

The upper Dini derivative of a function $h : (a,b) \to R, -\infty \leq a < b \leq +\infty$ is defined as

$$D^+ h(t) = \limsup_{s \to 0^+} \frac{h(t+s) - h(t)}{s}.$$

Suppose h is continuous on (a,b). Then h is non-increasing on (a,b) if and only if $D^+ h(t) \leq 0$ for any $t \in (a,b)$ (see [11] for details). The next result is given for the calculation of Dini derivative [4], [21].

Lemma 3.1: Let $V_i(t,x) : R \times R^d \to R\ (i = 1, \cdots, n)$ be C^1 and $V(t,x) = \max_{i=1,\cdots,n} V_i(t,x)$. If $\mathcal{I}(t) = \{i \in \{1, 2, \cdots, n\} : V(t,x(t)) = V_i(t,x(t))\}$ is the set of indices where the maximum is reached at t, then $D^+ V(t,x(t)) = \max_{i \in \mathcal{I}(t)} \dot{V}_i(t,x(t))$.

The following lemma was obtained in [18], which is also useful in what follows.

Lemma 3.2: Suppose $K \subset R^d$ is a convex set and $x_a, x_b \in R^d$. Then

$$\langle x_a - \mathcal{P}_K(x_a), x_b - x_a \rangle \leq |x_a|_K \cdot \left| |x_a|_K - |x_b|_K \right|. \tag{13}$$

Particularly, if $|x_a|_K > |x_b|_K$, then

$$\langle x_a - \mathcal{P}_K(x_a), x_b - x_a \rangle \leq -|x_a|_K \cdot (|x_a|_K - |x_b|_K). \tag{14}$$

Then we prove the following lemma.

Lemma 3.3: $D^+ \bar{d}(t) \leq 0$ for any $t \geq 0$.

Proof: According to (7), one has

$$\begin{aligned} \frac{dh_i(t)}{dt} &= 2\langle x_i - \mathcal{P}_{\mathcal{S}_0}(x_i), \dot{x}_i \rangle \\ &= 2\langle x_i - \mathcal{P}_{\mathcal{S}_0}(x_i), \sum_{j \in N_i(\sigma(t))} a_{ij}(x,t)(x_j - x_i) \\ &\quad + \mathcal{P}_{\mathcal{S}_i}(x_i) - x_i. \rangle \end{aligned} \tag{15}$$

Then, based on Lemma 3.1 and let $\mathcal{I}(t)$ denote the set containing all the agents that reach the maximum of $\bar{d}(t)$ at time t, we obtain

$$\begin{aligned} D^+\bar{d}(t) &= \max_{i\in\mathcal{I}(t)} \frac{d}{dt}d_i(t) \\ &= 2\max_{i\in\mathcal{I}(t)}[\langle x_i-\mathcal{P}_{\mathcal{S}_0}(x_i), \sum_{j\in N_i(\sigma(t))} a_{ij}(x_j-x_i) \\ &\quad +\mathcal{P}_{\mathcal{S}_i}(x_i)-x_i\rangle]. \end{aligned} \tag{16}$$

Furthermore, for any $i\in\mathcal{I}(t)$, according to (14) of Lemma 3.2, one has

$$\langle x_i-\mathcal{P}_{\mathcal{S}_0}(x_i), x_j-x_i\rangle \leq 0 \tag{17}$$

for any $j\in L_i(\sigma(t))$ since it always holds that $|x_j|_{\mathcal{S}_0}\leq |x_i|_{\mathcal{S}_0}$. Moreover, it is easy to see that for any $i\in\mathcal{V}$,

$$\begin{aligned} \langle x_i-\mathcal{P}_{\mathcal{S}_0}(x_i), \mathcal{P}_{\mathcal{S}_i}(x_i)-x_i\rangle &= \langle (x_i-\mathcal{P}_{\mathcal{S}_i}(x_i)) \\ +(\mathcal{P}_{\mathcal{S}_i}(x_i)-\mathcal{P}_{\mathcal{S}_0}(x_i)), \mathcal{P}_{\mathcal{S}_i}(x_i)-x_i\rangle. \end{aligned} \tag{18}$$

Next, in light of (6), we obtain

$$\langle \mathcal{P}_{\mathcal{S}_i}(x_i)-\mathcal{P}_{\mathcal{S}_0}(x_i), \mathcal{P}_{\mathcal{S}_i}(x_i)-x_i\rangle\leq 0 \tag{19}$$

since we always have $\mathcal{P}_{\mathcal{S}_0}(x_i)\in\mathcal{S}_i$ for all $i=1,\cdots,N$.

Therefore, with (16), (18) and (19), one has

$$\begin{aligned} D^+\bar{d}(t) &= \max_{i\in\mathcal{I}(t)} \frac{d}{dt}d_i(t) \\ &\leq 2\max_{i\in\mathcal{I}(t)}\langle x_i-\mathcal{P}_{\mathcal{S}_i}(x_i), \mathcal{P}_{\mathcal{S}_i}(x_i)-x_i\rangle \\ &\leq 2\max_{i\in\mathcal{I}(t)}[-|x_i|^2_{\mathcal{S}_i}] \\ &\leq 0 \end{aligned} \tag{20}$$

which leads to the conclusion. □

With Lemma 3.3, there exists a constant $\bar{d}^*\geq 0$ such that $\lim_{t\to\infty}\bar{d}(t)=\bar{d}^*$. Clearly the optimal solution set convergence will be achieved for system (9) if and only if $\bar{d}^*=0$.

Furthermore, since it always holds that $d_i(t)\leq\bar{d}(t)$, there exist constants $0\leq\theta_i\leq\eta_i\leq\bar{d}^*$ such that

$$\liminf_{t\to\infty} d_i(t)=\theta_i, \quad \limsup_{t\to\infty} d_i(t)=\eta_i,$$

for all $i=1,\cdots,N$.

Then we consider the following system:

$$\dot{x}_i = \sum_{j\in N_i(\sigma(t))} a_{ij}(x,t)(x_j-x_i)+\delta_i(t),\ i=1,\cdots,N \tag{21}$$

where $\delta_i(t):R_{\geq 0}\to R, i=1,\cdots,N$. The following conclusion holds.

Proposition 3.1: Suppose $\lim_{t\to\infty}\delta_i(t)=0, i=1,\cdots,N$. Then system (21) achieves the global consensus if $\mathcal{G}_{\sigma(t)}$ is UJSC.

Proof: Let

$$\hbar(t)=\max_{i\in\mathcal{V}}\{x_i(t)\},\quad \ell(t)=\min_{i\in\mathcal{V}}\{x_i(t)\}$$

be the maximum and minimum state value at time t. Denote $\mathcal{H}(x(t))=\hbar(t)-\ell(t)$.

Then since $\lim_{t\to\infty}\delta_i(t)=0$, we have that $\forall\varepsilon>0, \exists\hat{T}(\varepsilon)>0$ such that $|\delta_i(t)|<\varepsilon, t>\hat{T}$. Take $k_0\in\mathcal{V}$ with $x_{k_0}(sK_0)=\ell(sK_0)$, where $K_0=(N-1)T, s=0,1,\ldots$. Then it is not hard to find that for all $t\in[sK_0,(s+1)K_0]$,

$$x_{k_0}(t) \leq \alpha_0\ell(sK_0)+(1-\alpha_0)\hbar(sK_0)+\theta_0\varepsilon.$$

where $\alpha_0\triangleq\frac{1}{2}e^{-(N-1)a^*K_0}$ and $\theta_0\triangleq K_0+\frac{1}{(N-1)a^*}$.

Furthermore, since $\mathcal{G}_{\sigma(t)}$ is UJSC, similar estimations can be carried out on k_0's neighbors, neighbors' neighbors, and so on. Then we can find two constants $0<\alpha_{N-1}<1$ and $\gamma_0>0$ to ensure the following inequality:

$$\mathcal{H}(x((s+1)K_0))\leq(1-\alpha_{N-1})\mathcal{H}(x(sK_0))+\gamma_0\varepsilon. \tag{22}$$

Since s can be any nonnegative integer in (22), the conclusion follows immediately.

IV. Set Convergence

In this section we give a result for set convergence and then prove Theorem 3.1.

At first we give another proposition.

Proposition 4.1: Suppose $\mathcal{G}_{\sigma(t)}$ is UJSC. If $\theta_i=\eta_i=\bar{d}^*$ for all $i=1,\cdots,N$, then $\bar{d}^*=0$.

Proof: Based on the definitions of θ_i and η_i, one has

$$\lim_{t\to+\infty} d_i(t)=\bar{d}^*,\quad i=1,\cdots,N$$

when $\theta_i=\eta_i=\bar{d}^*$ holds for for all $i=1,\cdots,N$. Thus, for any $\varepsilon>0$, there exists $T_1(\varepsilon)>0$ such that, when $t>T_1(\varepsilon)$,

$$d_i(t)\in(\bar{d}^*-\varepsilon,\bar{d}^*+\varepsilon),\quad i=1,\cdots,N. \tag{23}$$

We will prove $\bar{d}^*=0$ by contradiction. Suppose $\bar{d}^*>0$ in the following.

First we have the following claim.

Claim. $\lim_{t\to+\infty}|x_i|_{\mathcal{S}_i}=0$ for all $i=1,\cdots,N$.

According to (15), (18) and (19), we obtain

$$\begin{aligned} \frac{dh_i(t)}{dt} &\leq -2|x_i|^2_{\mathcal{S}_i}+2\langle x_i-\mathcal{P}_{\mathcal{S}_0}(x_i), \\ &\quad \sum_{j\in N_i(\sigma(t))} a_{ij}(x,t)(x_j-x_i)\rangle. \end{aligned} \tag{24}$$

Furthermore, according to Lemma 3.2 and (23), one has that for any $\varepsilon>0$, there exists $T_2(\varepsilon)>0$ such that, when $t>T_2(\varepsilon)$,

$$\langle x_i-\mathcal{P}_{\mathcal{S}_0}(x_i), x_j-x_i\rangle\leq|x_i|_{\mathcal{S}_0}\cdot||x_i|_{\mathcal{S}_0}-|x_j|_{\mathcal{S}_0}|\leq\varepsilon \tag{25}$$

for all $i\in\mathcal{V}$ and $j\in N_i(\sigma(t))$. Thus, if it does not holds that $\lim_{t\to+\infty}|x_i|_{\mathcal{S}_i}=0$ for all $i=1,\cdots,N$, there exist a node i_0 and two constant $\tau_0, M_0>0$ such that

$$|x_{i_0}(t)|_{\mathcal{S}_{i_0}}\in[\frac{M_0}{2},M_0],\quad t\in[t_k,t_k+\tau_0] \tag{26}$$

for a time serial

$$0<t_1<\cdots<t_k<t_{k+1}<\cdots$$

with $t_k+\tau_0\leq t_{k+1}$ for $k=1,2,\cdots$. With (24), (25) and (26), it follows that, for any $\varepsilon>0$, when $t_k>\max\{T_1,T_2\}$, one has

$$\frac{dh_{i_0}(t)}{dt}\leq-\frac{1}{2}M_0^2+\varepsilon,\quad t\in[t_k,t_k+\tau_0]. \tag{27}$$

Note that (27) contradicts (23), and then the claim is proved.

Therefore, for any $\varepsilon > 0$, there exists $T_3(\varepsilon) > 0$ such that when $t > T_3$,

$$d_i(t) = |x_i(t)|^2_{\mathcal{S}_0} \in (\bar{d}^* - \varepsilon, \bar{d}^* + \varepsilon), \quad i = 1, \cdots, N. \tag{28}$$

and

$$|x_i(t)|_{\mathcal{S}_i} < \varepsilon, \quad i = 1, \cdots, N. \tag{29}$$

Then, based on Proposition 3.1 and (29), when $\mathcal{G}_{\sigma(t)}$ is UJSC, one has

$$\lim_{t \to +\infty} x_i(t) - x_j(t) = 0, \quad i, j = 1, \cdots, N,$$

which implies that for any $\varepsilon > 0$, there exists $T_4(\varepsilon) > 0$ such that when $t > T_4$,

$$|x_i(t) - x_j(t)| < \varepsilon, \quad i, j = 1, \cdots, N. \tag{30}$$

With (29) and (30), for any $\varepsilon > 0$, when $t > \max\{T_3, T_4\}$, one has that

$$|x_i(t)|_{\mathcal{S}_j} < 2\varepsilon, \quad i, j = 1, \cdots, N, \tag{31}$$

which implies

$$|x_i(t)|_{\mathcal{S}_0} < 2\varepsilon, \quad i, j = 1, \cdots, N. \tag{32}$$

Thus, (32) contradicts (28) when ε is sufficiently small. Therefore, $\bar{d}^* > 0$ does not hold and the conclusion holds immediately. $\square$

Then we have the following result on optimal set convergence.

Theorem 4.1: System (9) achieves the optimal solution set convergence if $\mathcal{G}_{\sigma(t)}$ is UJSC.

Proof: We also prove the conclusion by contradiction. Suppose $\bar{d}^* > 0$.

Then, for any $\varepsilon > 0$, there exists $T_1(\varepsilon) > 0$ such that, when $t > T_1(\varepsilon)$,

$$d_i(t) \in (0, \bar{d}^* + \varepsilon), \quad i = 1, \cdots, N. \tag{33}$$

According to Proposition 4.1, there exist at least one agent $i_0 \in \mathcal{V}$ such that $0 \leq \theta_{i_0} < \eta_{i_0} \leq \bar{d}^*$. Take $\zeta_0 = \sqrt{\frac{1}{2}(\theta_{i_0} + \eta_{i_0})}$. Then there exists a time serial

$$0 < \hat{t}_1 < \cdots < \hat{t}_k < \cdots$$

with $\lim_{t\to\infty} \hat{t}_k = \infty$ such that $h_{i_0}(\hat{t}_k) = \zeta_0^2$ for all $k = 1, 2, \cdots$.

Furthermore, take $\hat{t}_{k_0} > T_1$, and according to (24) and Lemma 3.2, one has for all $t > \hat{t}_{k_0}$,

$$\begin{aligned}\frac{dh_{i_0}(t)}{dt} &\leq 2 \sum_{j \in N_i(\sigma(t))} a_{i_0 j}(x,t) \langle x_{i_0} - \mathcal{P}_{\mathcal{S}_0}(x_{i_0}), x_j - x_{i_0} \rangle \\ &\leq 2(N-1)a^* |x_{i_0}(t)|_{\mathcal{S}_0} (\sqrt{\bar{d}^* + \varepsilon} - |x_{i_0}(t)|_{\mathcal{S}_0}),\end{aligned}$$

which is equivalent to

$$\frac{d\sqrt{h_{i_0}(t)}}{dt} \leq -(N-1)a^* \sqrt{h_{i_0}(t)} + (N-1)a^* \sqrt{\bar{d}^* + \varepsilon}.$$

As a result, for $t \in (\hat{t}_{k_0}, \infty)$, we have

$$\begin{aligned}\sqrt{h_{i_0}(t)} &\leq e^{-(N-1)a^*(t - \hat{t}_{k_0})} \sqrt{h_{i_0}(\hat{t}_{k_0})} + \\ &\quad (1 - e^{(N-1)a^*(t - \hat{t}_{k_0})}) \sqrt{\bar{d}^* + \varepsilon} \\ &\leq e^{-(N-1)a^*(t - \hat{t}_{k_0})} \zeta_0 \\ &\quad + (1 - e^{(N-1)a^*(t - \hat{t}_{k_0})}) \sqrt{\bar{d}^* + \varepsilon}. \end{aligned} \tag{34}$$

Next, since $\mathcal{G}_{\sigma(t)}$ is uniformly jointly strongly connected, there is at lest one arc leaving from i_0 entering $i_1 \in \mathcal{V}$ in $\mathcal{G}([\hat{t}_{k_0}, \hat{t}_{k_0} + T))$. Moreover, it is not hard to see that this arc exits for at least τ_D during $t \in [\hat{t}_{k_0}, \hat{t}_{k_0} + T + 2\tau_D)$, which implies that $(i_0, i_1) \in \mathcal{G}_{\sigma(t)}$ for some $t \in [\tilde{t}_1, \tilde{t}_1 + \tau_D) \subseteq [\hat{t}_{k_0}, \hat{t}_{k_0} + T + 2\tau_D)$. Denote $T_0 = T + 2\tau_D$. Then, one has

$$\begin{aligned}\sqrt{h_{i_0}(t)} &\leq -e^{(N-1)a^* T_0} \zeta_0 + (1 - e^{(N-1)a^* T_0}) \sqrt{\bar{d}^* + \varepsilon} \\ &\doteq \xi_1 \end{aligned} \tag{35}$$

for all $t \in (\hat{t}_{k_0}, \hat{t}_{k_0} + T_0)$. Thus, for $t \in [\tilde{t}_1, \tilde{t}_1 + \tau_D)$, one has

$$\begin{aligned}\frac{dh_{i_1}(t)}{dt} &\leq 2 \sum_{j \in N_i(\sigma(t)) \setminus i_0} a_{i_1 j}(x,t) \langle x_{i_1} - \mathcal{P}_{\mathcal{S}_0}(x_{i_1}), \\ &\qquad x_j - x_{i_1} \rangle \\ &\quad + a_{i_1 i_0}(x,t) \langle x_{i_1} - \mathcal{P}_{\mathcal{S}_0}(x_{i_1}), x_{i_0} - x_{i_1} \rangle \\ &\leq 2(N-2)a^* |x_{i_1}(t)|_{\mathcal{S}_0} (\sqrt{\bar{d}^* + \varepsilon} - |x_{i_1}(t)|_{\mathcal{S}_0}) \\ &\quad - a_* |x_{i_1}(t)|_{\mathcal{S}_0} (|x_{i_1}(t)|_{\mathcal{S}_0} - \xi_1), \end{aligned} \tag{36}$$

which is equivalent to

$$\begin{aligned}\frac{d\sqrt{h_{i_1}(t)}}{dt} &\leq -[(N-2)a^* + a_*] \sqrt{h_{i_1}(t)} \\ &\quad + (N-2)a^* \sqrt{\bar{d}^* + \varepsilon} + a_* \xi_1. \end{aligned} \tag{37}$$

Then we obtain

$$\begin{aligned}\sqrt{h_{i_1}(t)} \leq{}& e^{-[(N-2)a^* + a_*](t - \tilde{t}_1)} \sqrt{h_{i_1}(\tilde{t}_1)} \\ &+ (1 - e^{-[(N-2)a^* + a_*](t - \tilde{t}_1)}) \\ &\cdot \frac{(N-2)a^* \sqrt{\bar{d}^* + \varepsilon} + a_* \xi_1}{(N-2)a^* + a_*}\end{aligned}$$

for for $t \in [\tilde{t}_1, \tilde{t}_1 + \tau_D)$, which leads to

$$\begin{aligned}\sqrt{h_{i_1}(\tilde{t}_1 + \tau_D)} &\leq e^{-[(N-2)a^* + a_*]\tau_D} \sqrt{\bar{d}^* + \varepsilon} \\ &\quad + (1 - e^{-[(N-2)a^* + a_*]\tau_D}) \\ &\quad \times \frac{(N-2)a^* \sqrt{\bar{d}^* + \varepsilon} + a_* \xi_1}{(N-2)a^* + a_*} \\ &\doteq \zeta_1 \end{aligned} \tag{38}$$

Therefore, based on similar analysis for (34), one has when $t \in [\tilde{t}_1 + \tau_D, \infty)$

$$\begin{aligned}\sqrt{h_{i_1}(t)} &\leq e^{-(N-1)a^*(t - (\tilde{t}_1 + \tau_D))} \zeta_1 \\ &\quad + (1 - e^{-(N-1)a^*(t - (\tilde{t}_1 + \tau_D))}) \sqrt{\bar{d}^* + \varepsilon}. \end{aligned} \tag{39}$$

Note that we have that $\zeta_0 < \zeta_1 < \bar{d}^*$. Therefore, we can proceed similar analysis on time intervals $(\hat{t}_{k_0}+T_0, \hat{t}_{k_0}+2T_0), (\hat{t}_{k_0}+2T_0, \hat{t}_{k_0}+3T_0), \cdots, (\hat{t}_{k_0}+(N-1)T_0, \hat{t}_{k_0}+NT_0)$ respectively, and then get similar estimation as (34) and (39) by $\zeta_0 < \zeta_1 < \cdots < \zeta_{N-1} < \bar{d}^*$ for agents $i_2, \cdots, i_{N-1}$ with $\mathcal{V} = \{i_0, i_1, \cdots, i_{N-1}\}$. Thus, we obtain

$$\begin{aligned} \sqrt{h_{i_j}(\hat{t}_{k_0}+NT_0))} \quad \leq \quad & e^{-(N-1)T_0 a^*}\zeta_{N-1} \\ & +(1-e^{-(N-1)NT_0 a^*}) \times \sqrt{\bar{d}^*+\varepsilon} \end{aligned}$$

for all $j = 0, 1, \cdots, N-1$, which contradicts the definition of $\bar{d}^*$ since

$$e^{-(N-1)T_0 a^*}\zeta_{N-1} + (1-e^{-(N-1)NT_0 a^*})\sqrt{\bar{d}^*+\varepsilon} < \sqrt{\bar{d}^*}$$

for sufficiently small ε. This completes the proof. □

Then we prove the main result:

Proof of Theorem 3.1: In fact, it is not hard to see that the conclusion hold by combining Proposition 3.1 and Theorem 4.1. □

Remark 4.1: UJSC is sufficient, but not necessary to guarantee an optimal consensus for System 9. However, simple examples can be constructed to show that weaker requirement for connectedness, such as, uniformly jointly quasi-strongly connectivity (UQSC) is not enough for optimal consensus, although it has been shown that UQSC can ensure a consensus for nonlinear multi-agent systems [21].

V. Conclusions

This paper addressed an optimal consensus problem for multi-agent systems. With time-varying interconnection topologies and uniformly joint connectivity assumption, the considered multi-agent system achieved not only a consensus, but also an optimal one by agreeing within the global solution set of a sum of objective functions corresponding to multiple agents. Moreover, control laws applied to the agents were nonlinear and distributed.

References

[1] J. Aubin and A. Cellina. *Differential Inclusions*. Berlin: Speringer-Verlag, 1984

[2] R. T. Rockafellar. *Convex Analysis.* New Jersey: Princeton University Press, 1972.

[3] C. Godsil and G. Royle. *Algebraic Graph Theory.* New York: Springer-Verlag, 2001.

[4] J. Danskin. The theory of max-min, with applications, *SIAM J. Appl. Math.*, vol. 14, 641-664, 1966.

[5] C. Berge and A. Ghouila-Houri. *Programming, Games, and Transportation Networks*, John Wiley and Sons, New York, 1965.

[6] D. Cheng, J. Wang, and X. Hu, An extension of LaSalle's invariance principle and its applciation to multi-agents consensus, *IEEE Trans. Automatic Control*, 53, 1765-1770, 2008.

[7] F. Clarke, Yu.S. Ledyaev, R. Stern, and P. Wolenski, *Nonsmooth Analysis and Control Theory*. Speringer-Verlag, 1998

[8] S. Martinez, J. Cortés, and F. Bullo. Motion coordination with distributed information, *IEEE Control Systems Magazine*, vol. 27, no. 4, 75-88, 2007.

[9] W. Ren and R. Beard, *Distributed Consensus in Multi-vehicle Cooperative Control*, Springer-Verlag, London, 2008.

[10] R. Olfati-Saber, Flocking for multi-agent dynamic systems: algorithms and theory, *IEEE Trans. Automatic Control*, 51(3): 401-420, 2006.

[11] N. Rouche, P. Habets, and M. Laloy. *Stability Theory by Liapunov's Direct Method*, New York: Springer-Verlag, 1977.

[12] Y. Hong, L. Gao, D. Cheng, and J. Hu. Lyapuov-based approach to multi-agent systems with switching jointly connected interconnection. *IEEE Trans. Automatic Control*, vol. 52, 943-948, 2007.

[13] R. Olfati-Saber and R. Murray. Consensus problems in the networks of agents with switching topology and time dealys, *IEEE Trans. Automatic Control*, vol. 49, no. 9, 1520-1533, 2004.

[14] Y. Hong, J. Hu, and L. Gao. Tracking control for multi-agent consensus with an active leader and variable topology. *Automatica*, vol. 42, 1177-1182, 2006.

[15] H. G. Tanner, A. Jadbabaie, G. J. Pappas, Flocking in fixed and switching networks, *IEEE Trans. Automatic Control*, 52(5): 863-868, 2007.

[16] F. Xiao and L. Wang, State consensus for multi-agent systems with swtiching topologies and time-varying delays, *Int. J. Control*, 79, 10, 1277-1284, 2006.

[17] A. Jadbabaie, J. Lin, and A. S. Morse. Coordination of groups of mobile agents using nearest neighbor rules. *IEEE Trans. Automatic Control*, vol. 48, no. 6, 988-1001, 2003.

[18] G. Shi and Y. Hong, Global target aggregation and state agreement of nonlinear multi-agent systems with switching topologies, *Automatica*, vol. 45, 1165-1175, 2009.

[19] Y. Cao and W. Ren, Containment control with multiple stationary or dynamic leaders under a directed interaction graph, *Proc. of Joint 48th IEEE Conf. Decision & Control/28th Chinese Control Conference*, Shanghai, China, Dec. 2009, pp. 3014-3019.

[20] J. Tsitsiklis, D. Bertsekas, and M. Athans. Distributed asynchronous deterministic and stochastic gradient optimization algorithms, *IEEE Trans. Automatic Control*, 31, 803-812, 1986.

[21] Z. Lin, B. Francis, and M. Maggiore. State agreement for continuous-time coupled nonlinear systems. *SIAM J. Control Optim.*, vol. 46, no. 1, 288-307, 2007.

[22] Lin Wang and Lei Guo, Robust Consensus of Multi-Agent Systems under Directed Information Exchanges, *Chinese Control Conference*, 557 - 561, 2007

[23] L. Moreau, Stability of multiagent systems with time-dependent communication links, *IEEE Trans. Automatic Control*, 50, 169-182, 2005.

[24] S. Boyd and L. Vandenberghe, Convex Optimization. New York, NY: Cambridge University Press, 2004.

[25] A. Nedić, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis, Distributed subgradient methods and quantization effects, *in Proc. IEEE Conference on Decision and Control, Cancun, Mexico*, 2008, pp. 41774184.

[26] A. Nedić and D. P. Bertsekas, Incremental subgradient methods for nondifferentiable optimization, *SIAM Journal on Optimization*, vol. 12, no. 1, pp. 109138, 2001.

[27] A. Nedić and A. Ozdaglar, Distributed subgradient methods for multi-agent optimization, *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 4861, 2009.

[28] Nedić, A., Ozdaglar, A. & Parrilo, P. A.(2010). Constrained Consensus and Optimization in Multi-Agent Networks. *IEEE Transactions on Automatic Control*, vol. 55, no. 4, pp. 922-938.

[29] B. Johansson, M. Rabi, and M. Johansson, A simple peer-to-peer algorithm for distributed optimization in sensor networks, *in Proc. IEEE Conference on Decision and Control, New Orleans, LA,* 2007, pp. 47054710.

[30] B. Johansson, T. Keviczky, M. Johansson, and K. H. Johansson, Subgradient methods and consensus algorithms for solving convex optimization problems, *Proc. IEEE Conference on Decision and Control*, Cancun, Mexico, 2008, pp. 41854190.